

Project Proposal

ChatDB – Taking to Database Management Systems using Natural Language

I. Team Members Background and Skills

- Yu-Ching Huang (jhuang13@usc.edu)
 - Major: Applied Data Science
 - Background: Bachelor's in Business Administration major with Business Analytics and Finance concentrations and a Data Science minor
 - Skills: Python, SQL, Tableau, Java, Git
- Shih-Hui Huang (shihhuih@usc.edu)
 - Major: Applied Data Science
 - Background: Bachelor's and Master's in Economics
 - Skills: Python
- Yu-Chen Lu (ylu74747@usc.edu)
 - Major: Applied Data Science
 - Background: Bachelor's in Economics
 - Skills: Python

II. Project Requirements

In this project, we will develop a ChatDB system that allows users to interact with the both SQL and NoSQL databases, including exploring database schemas, retrieving data, modifying data, etc., using natural language.

III. Planned Implement

We plan to implement Llama 2 locally to transform natural language into SQL queries. For SQL databases, we will use MySQL, and for NoSQL databases, we will utilize Firebase and MongoDB.

The interface will support various functions in both SQL and NoSQL databases. Users would be able to explore database schemas and data by identifying tables or collections, viewing their attributes, and retrieving sample data. The ChatDB will be able to convert natural language into appropriate SQL or NoSQL commands to execute data retrieval operations, supporting SQL clauses like SELECT, WHERE, JOIN, GROUP BY, and functions in NoSQL databases such as find, aggregate, and \$lookup. Additionally, the interface will handle data modification requests, enabling users to insert, update, and delete data seamlessly using natural language instructions.

In addition, we will develop a web-browser-based user interface that will provide an intuitive and user-friendly platform for users to interact with ChatDB, allowing them to view responses in a structured format.

IV. Team Responsibilities (Subject to change)

- **Yu-Ching Huang:** SQL Database Implementation / Testing & Debugging
- **Shih-Hui Huang:** NoSQL Database Implementation
- **Yu-Chen Lu:** Research for LLM Options / Web-browser based UI Development

V. Timeline (Subject to change)

[illegible]